ЗК

УДК 519.7

И.В. Каменева
Харьковский национальный университет радиоэлектроники, г. Харьков
Украина, , 61166, г. Харьков, просп. Ленина, 14

Преобразование словоформы с помощью метода расслоения предиката в морфологии русского языка

I.V. Kameneva
Kharkov National University of Radioelectronics, c. Kharkov
Ukraine, , 61166, *c. Kharkov, Lenin ave., 14*

Word Form Transformation by Separation Finite Predicate Method in Russian Morphology

І.В. Камєнєва
Харківський національний університет радіоелектроніки, м. Харків
Україна, , 61166, м. Харків, просп. Леніна, 14

Перетворення словоформи за допомогою методу розшарування предиката в морфології російської мови

В статье рассматривается метод расслоения конечного предиката для морфологических структур русского языка, где исследуется признак контекста и образование словоформы для морфологической задачи моделирования процесса склонения полных непритяжательных имен прилагательных, представленных в звуковой форме.
Ключевые слова: расслоение конечного предиката, реляционная сеть, бинаризация.

In the article, separation finite predicate method for morphological structures of the Russian language is considered. The attribute context and word form creation for morphological task of process modeling an adjectives presented in sound form are investigated.
Key Words: the separation finite predicate method, relation network, binarization.

У роботі розглядається метод розшарування кінцевого предиката для морфологічних структур російської мови, де досліджується ознака контексту та утворення словоформи для морфологічної задачі моделювання процесу відмінювання повних неприсвійних прикметників, представлених в звуковій формі.
Ключові слова: розшарування кінцевого предиката, реляційна мережа, бінаризація.

Введение

Для создания модели языка необходимо формализовать разнообразные и многочисленные функции, реализуемые человеком при обработке речевой и текстовой информации. Важную роль при разработке математического обеспечения информационных систем играют реляционные и логические способы представления знаний. Одним из эффективных универсальных математических средств для описания конечных информационных массивов является алгебра конечных предикатов, которая выражает предикаты заданными отношениями и отношения – предикатами.

Отношения выражают свойства предметов и связи между ними. Они представляют собой универсальное средство описания любых объектов. Язык человека, представляющий собой главное средство духовного общения людей, можно рассматривать

лишь как механизм для выражения отношений. Обмен мыслями между людьми осуществляется только за счет передачи и приема отношений. Каждая мысль представляет собой какое-то отношение. Имеется множество бинарных отношений. Это множество представляет собой систему, поскольку является декомпозицией одного отношения. Вся эта система может быть представлена в виде неориентированного графа и называется логической реляционной сетью. Ребра этого графа называются ветвями сети, верши ны – полюсами. Ветви представляют собой бинарные отношения. Полюса сети – это ячейки памяти, в каждой из которых содержится множество значений некоторой предметной переменной. Множество всех значений переменной – домен, любое его подмножество будет унарным отношением; это подмножество характеризует знание о значениях данной предметной переменной, которые содержатся в данном полюсе. [1]

Логическая реляционная сеть предназначена для решения системы уравнений, задаваемых соответствующей моделью. Аналитически можно решить любую систему уравнений алгебры предикатов, характеризующую данную модель. При этом можно задавать любое знание о значениях любых предметных переменных и получать знания о значениях любых других предметных переменных. Аналитический метод работает безотказно.

Логическая реляционная сеть работает с использованием потоков, каждый поток привязывается к определенному полюсу и пересекает все связанные с ним ветвями полюса на основании таблицы отношений соответствующей ветви. В процессе пересечения создается множество пересечения на основании таблицы отношений ветви и состояния, связанного с потоком полюса, после чего проводится логическое умножение между множеством пересечения и множеством состояния пересекаемого полюса, и результат записывается в состояние полюса. В процессе решения происходит синхронизация потоков, запускаются только те потоки, состояние полюсов которых изменилось после последнего запуска потоков. Если состояние ни одного из полюса потоков уже не изменяется, все потоки останавливаются.[2], [3]

Перед нами стоит задача проверить и оценить работоспособность метода расслоения предиката для полной модели склонения непритяжательных имен прилагательных в звуковой форме [4], [5].

В данной статье рассматривается применение метода расслоения для признаков контекста и образования словоформы. На рис. 1 представлена реляционная сеть склонения полных непритяжательных имен прилагательных русского языка, представленных в звуковой форме, где выделяется отдельная часть для того, чтобы определить насколько эффективно применяется метод расслоения предиката для образования словофрмы звуковой модели.

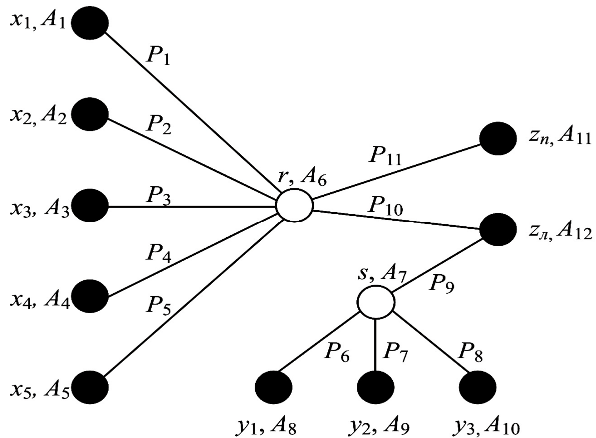


Рисунок 1 – Строение модели реляционной сети склонения полных непритяжательных имен прилагательных русского языка, представленных в звуковой форме

Метод расслоения предиката для образования словоформы

Нами успешно продемонстрирована работа с сетью, которая ранее была построена с помощью метода расслоения предиката в виде многополюсника для признаков слова, таких как y_1 – последний звук основы слова, y_2 – признак ударности-безударности, y_3 – мягкости-твердости, которые переходят соответственно в номер типа склонения. В результате были получены предикаты $F_1, F_2, F_3; G_1, G_2, G_3$ [6]. Продолжим использование метода расслоения предиката для перехода от номера типа склонения к левой части окончания.

Итак, остановимся на том, что пошагово выполним все этапы метода расслоения предиката для образования словоформы.

Допустим, что табл. 1 представлена левой частью окончания $z_л$ относительно типа склонения слова t со значениями 1-7.

Таблица 1 – $z_л$ относительно типа склонения слова t

t / $z_л$	1	2	3	4	5	6	7
ы	1	1	0	0	0	0	0
э	0	1	0	0	0	0	0
о	1	0	0	1	0	0	0
и	0	0	1	1	0	0	0
а,у	1	1	0	1	0	0	0
е,я,ю	0	0	1	0	0	0	0
о’а’у’	0	0	0	0	1	1	0
ы’	0	0	0	0	1	0	0
и’	0	0	0	0	0	1	1
ё, я’,ю’	0	0	0	0	0	0	1

СДНФ исходного предиката *P* равна:

w	1	2	3	4	5	6	7
$z_л$	ы,о,а,у	ы,э,а,у	и,е,я,ю	и,о,а,у	ы’,о,’а’,у’	и’,о’,а’,у’	и’,ё, я’,ю’

$$P(z_л, t) = z_л^{ы}(t^1 \vee t^2) \vee (z_л^{э} t^2 \vee z_л^{о})(t^1 \vee t^4) \vee z_л^{и}(t^3 \vee t^4) \vee (z_л^{а} \vee z_л^{у})(t^1 \vee t^2 \vee t^4) \vee$$
$$\vee (z_л^{е} \vee z_л^{я} \vee z_л^{ю}) t^3 \vee (z_л^{о'} \vee z_л^{а'} \vee z_л^{у'})(t^5 \vee t^6) \vee z_л^{ы'} t^6 \vee z_л^{и'}(t^6 \vee t^7) \vee (z_л^{ё} \vee z_л^{я} \vee z_л^{ю}) t^7$$

Вводим предикат эквивалентности и получаем классификаторы предиката $P(z_л, w)$, где w – класс в явном виде, в неявном виде $F_i(z_л, w)$.

$$F_4(z_л, w) = (z_л^ы \vee z_л^о \vee z_л^а \vee z_л^у)w^1 \vee (z_л^ы \vee z_л^э \vee z_л^а \vee z_л^у)w^2 \vee z_л^о)(w^1 \vee w^4) \vee$$
$$(z_л^и \vee z_л^е \vee z_л^я \vee z_л^ю)w^3 \vee (z_л^ы \vee z_л^о \vee z_л^а \vee z_л^у)w^4 \vee (z_л^{ы'} \vee z_л^{о'} \vee z_л^{а'} \vee z_л^{у'})w^5 \vee$$
$$\vee (z_л^{и'} \vee z_л^{о'} \vee z_л^{а'} \vee z_л^{у'})w^6 \vee (z_л^{и'} \vee z_л^ё \vee z_л^{я'} \vee z_л^{ю'})w^7.$$

Перешли в неявном виде к ячейкам таблицы со значениями типов склонения.

Таблица 2 – h_1 относительно признаков y_2 – ударности-безударности, y_3 – мягкости-твердости

	y_2, y_3			
h_1	У Т	У М	Б Т	Б М
б	{ы, о, а, у}	{и, е, я, ю}	{ы′, о′, а′, у′}	{и′, ё, я′, ю′}
г	–	{и, е, я, ю}	–	{и′, ё, я′, ю′}
ч	–	{и, о, а, у}	{и′, о′, а′, у′}	–
ж	{ы, э, а, у}	–	{ы′, о′, а′, у′}	–
	w			

Табл. 2 по своим значениям совпадает с таблицей зависимости признаков слова h_1 y_2 y_3 от типа склонения w, полученной для сети характеризующей признаки слова.

Таблица 3 – Зависимость y_1 от типа склонения t

	$y_2 y_3$			
y_1	У Т	У М	Б Т	Б М
б	1	3	5	7
ч	–	3	–	7
г	–	4	6	–
ж	2	–	5	–
	t			

В табл. 2 классы $z_л$ полностью совпадают со значениями типа склонения в табл. 3, следовательно, можно провести тот же переход от признаков контекста к номеру типа склонения (табл. 4), связав его с $z_л$, и получить предикат $R(h_1, z_л, w, s)$.

Таблица 4 – Переход от признаков контекста к номеру типа склонения

$$R(h_1, z_л, w, s) = h^б(z_л^ы \vee z_л^о \vee z_л^а \vee z_л^у)w^1 s^1 \vee (h^ч \vee h^б)(z_л^и \vee z_л^е \vee z_л^я \vee z_л^ю)w^3(s^2 \vee s^6) \vee$$
$$\vee h^г(z_л^и \vee z_л^о \vee z_л^а \vee z_л^у)w^4 s^3 \vee h^ж(z_л^ы \vee z_л^э \vee z_л^а \vee z_л^у)w^2 s^4 \vee (h^ч \vee h^б)(z_л^{и'} \vee z_л^{е'} \vee z_л^{я'} \vee$$
$$\vee z_л^{ю'})w^7(s^6 \vee s^{10}) \vee (h^ч \vee h^б)(z_л^{и'} \vee z_л^{е'} \vee z_л^{я'} \vee z_л^{ю'})w^7(s^6 \vee s^{10}) \vee (h^г \vee h^б)(z_л^{ы'} \vee z_л^{о'} \vee$$
$$\vee z_л^{а'} \vee z_л^{у'})w^5(s^7 \vee s^9) \vee h^ж(z_л^{и'} \vee z_л^{о'} \vee z_л^{а'} \vee z_л^{у'})w^6 s^8$$

Производим бинаризацию предиката R и находим предикат $G_4(z_л, w)$:

$$G_4(z_л, w) = (z_л^ы \vee z_л^о \vee z_л^а \vee z_л^у)w^1 \vee (z_л^и \vee z_л^е \vee z_л^я \vee z_л^ю)w^3 \vee$$
$$\vee (z_л^и \vee z_л^о \vee z_л^а \vee z_л^у)w^4 \vee (z_л^ы \vee z_л^э \vee z_л^а \vee z_л^у)w^2 \vee (z_л^{и'} \vee z_л^{е'} \vee z_л^{я'} \vee$$
$$\vee z_л^{ю'})w^7 \vee (z_л^{ы'} \vee z_л^{о'} \vee z_л^{а'} \vee z_л^{у'})w^5 \vee (z_л^{и'} \vee z_л^{о'} \vee z_л^{а'} \vee z_л^{у'})w^6$$

Строим ассоциативный слой переключательной сети $P(h_1, z_л, w, s) = g$ (рис. 2).

h_1	$z_л$	w	s
б	$\{ы, о, а, у\}$	1	1
ч	$\{и, е, я, ю\}$	3	2
г	$\{и, о, а, у\}$	4	3
ж	$\{ы, э, а, у\}$	2	4
б	$\{и, е, я, ю\}$	3	5
ч	$\{и', ё, я', ю'\}$	7	6
г	$\{ы', о', а', у'\}$	5	7
ж	$\{и, о, а, у\}$	6	8
б	$\{ы', о', а', у'\}$	5	9
б	$\{и', ё, я', ю'\}$	7	10

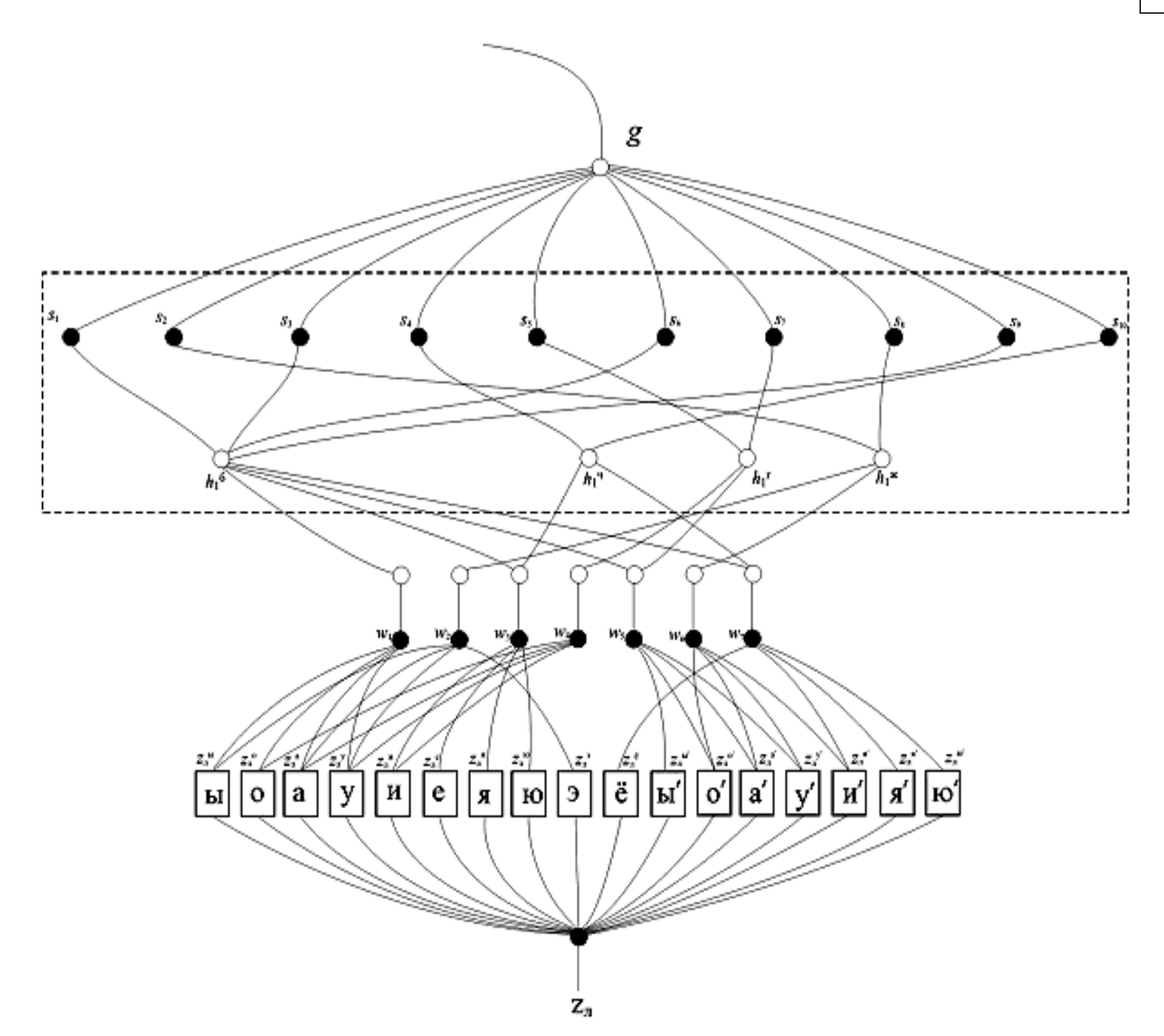


Рисунок 2 – Переключательная цепь

Задаем g = 1 и получаем реляционную сеть в виде многополюсника для отношения $P(h_1, z_л, w, s) = t$, построенную для признаков слова: y_1 – последний звук основы слова, y_2 – признак ударности-безударности, y_3 – мягкости-твердости, которая изображена на рис. 3 и расширена с добавлением признака словоформы $z_л$:

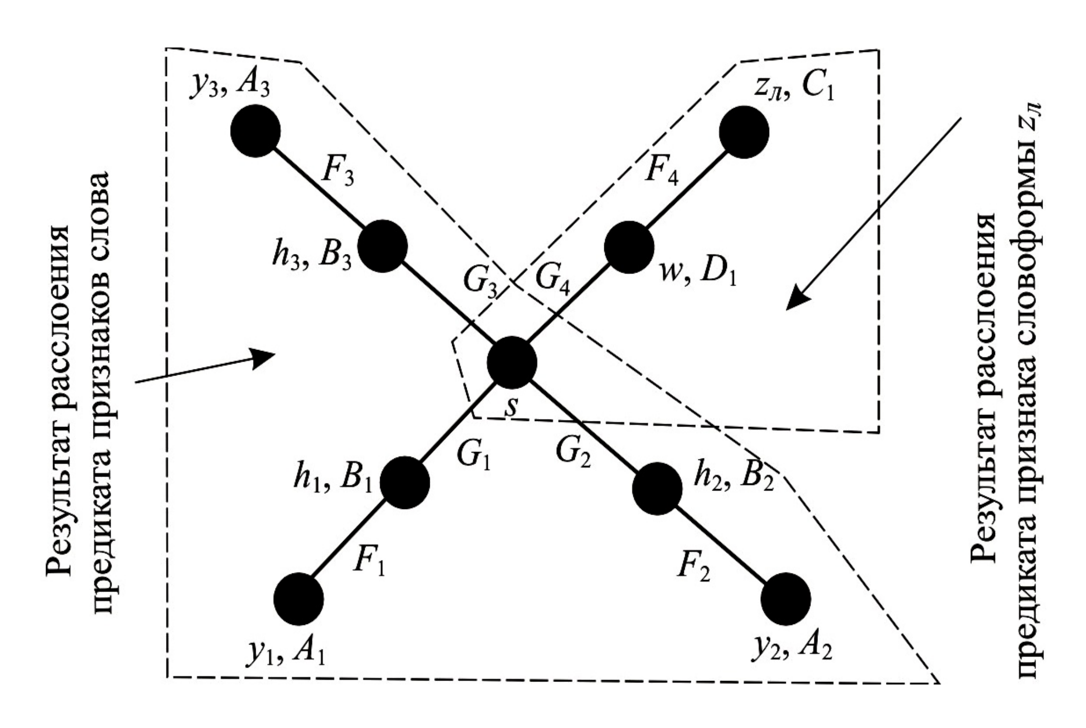


Рисунок 3 – Реляционная сеть

Выводы

С помощью метода расслоения предиката была усовершенствована модель склонения полных непритяжательных имен прилагательных звуковой модели, которая теперь дает возможность автоматизировать процесс перехода от конечного предиката к реляционной сети для морфологических структур. Логическая реляционная сеть была построена для признака контекста, признаков слова и образования словоформы. Как выяснилось, в представленной модели склонения признаки контекста минимизированы настолько, что применить метод расслоения предиката невозможно. В дальнейшем планируется доказать универсальность и эффективность метода расслоения предиката на различных морфологических структурах русского языка.

Литература

1. Бондаренко М.Ф. Мозгоподобные суктуры / М.Ф. Бондаренко, Ю.П. Шабанов-Кушнаренко. – Київ : Наукова думка.– 2011.– 160 с.
2. Бондаренко М.Ф. О реляционных сетях / М.Ф. Бондаренко, Н.П. Кругликова, И.А. Лещинская, Н.Е. Русакова, Ю.П. Шабанов-Кушнаренко // Бионика интеллекта. – 2010. – № 3. – С. 8-13.
3. Бондаренко М.Ф. О булевых реляционных сетях / М.Ф. Бондаренко, И.В. Каменева, Н.Е. Русакова, Ю.П. Шабанов-Кушнаренко, И.Ю. Шубин // Бионика интеллекта: научн.-техн. журнал. – 2011. – №1 (75). – С. 3 – 7.
4. Русакова Н. Е. Модель устной речи / Н. Е. Русакова // Бионика интеллекта: научн.-техн. журнал. – 2010. – №1(72). – С. 94 – 97.
5. Русакова Н.Е. О методе расслоения конечного предиката / Н.Е. Русакова // Бионика интеллекта: научн.-техн. журнал. – 2011. – №3 (77). – С. 50 – 53.
6. Каменева И.В., Русакова Н.Е. Использование метода расслоения предиката в морфологии русского языка / И. В. Каменева, Н.Е. Русакова // Системные технологии. – 2012. – № 4(81). С. 86 –94.

Literatura

1. Bondarenko M.F. Mozgopodobnye sruktury. Kyiv: Naukova dumka.2011. 160 s.
2. Bondarenko M.F. Bionika intellekta. 2010. № 3. S. 8-13.
3. Bondarenko M.F. Bionika intellekta: nauchn.-tehn. zhurnal. 2011. № 1 (75). S. 3-7.

4. Rusakova N. E. Bionika intellekta: nauchn.-tehn. zhurnal. 2010. № 1(72). S. 94-97.
5. Rusakova N.E. Bionika intellekta: nauchn.-tehn. zhurnal. 2011. № 3 (77). S. 50-53.
6. Kameneva I. V. Sistemnye tehnologii. 2012. № 4(81). S. 86-94.

RESUME

I.V. Kameneva

Word Form Transformation by Separation Finite Predicate Method in Russian Morphology

It is necessary to formalize diverse and multiple functions implemented in the processing of human voice and text information for model of language creation.

Relational and logical methods of knowledge representation play an important role in the software information systems development. The algebra of finite predicate is one of the effective universal mathematical tools for describing endpoint arrays that expresses predicates by relations and relations by predicates.

The task is to test and evaluate the performance of the separation finite predicate for an adjectives presented in a sound form.

In this paper, all steps of separation finite predicate method for morphological structures of the Russian language are implemented, i.e. equivalence of predicates are input, predicate classifiers are got, grading layer of switching circuit predicate are built.

The number of type of inducement was transferred to the left side of ending and transited from the attribute context to the number of type of inducement. Then binarization was performed, associative layer of switching network was built, and relational network was got.

The attribute context and word form creation for morphological task of process modeling of an adjectives presented in sound form were investigated.

Статья поступила в редакцию 01.06.2012.